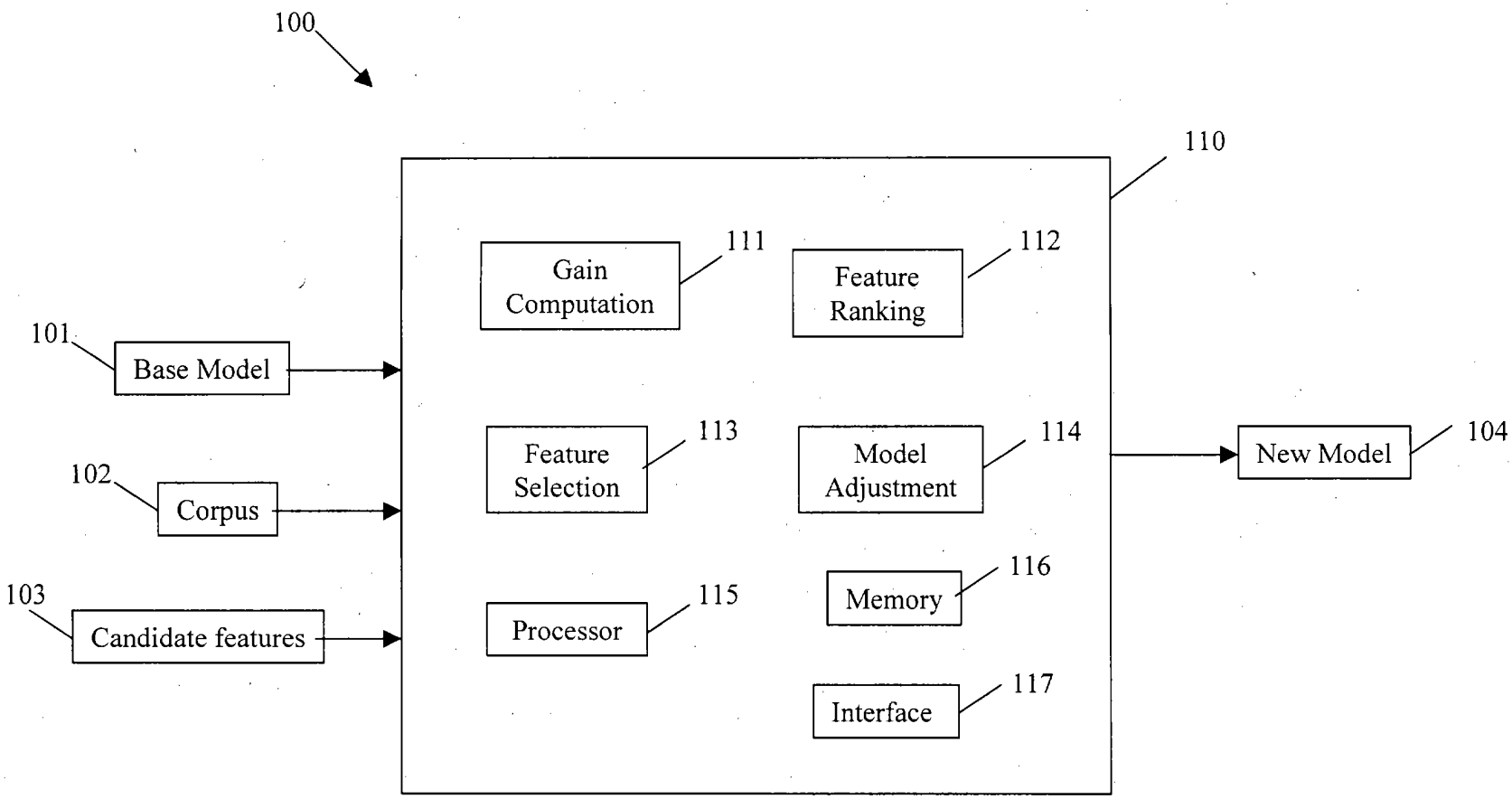
100
110
Gain Computation
111
Feature Ranking
112
101
Base Model
Feature Selection
113
Model Adjustment
114
New Model
104
102
Corpus
103
Candidate features
Processor
115
Memory
116
Interface
117

Figure 1

0. **Initialize**: $S = \varnothing$, $sum[1..I, 1..Y] = 1$,
$z[1..I] = Y$

1. **Gain computation:**
MaxGain = 0
for f in feature space F **do**
$\hat{\alpha} = \arg\max_{\alpha} G_{S \cup f}(\alpha)$
$\hat{g} = \max_{\alpha} G_{S \cup f}(\alpha)$
if MaxGain $< \hat{g}$ **then**
MaxGain = $\hat{g}$
$f^* = f$
$\alpha^* = \hat{\alpha}$

2. **Feature selection**:
$S = S \cup \{f^*\}$
$F = F - \{f^*\}$

3. **if** termination condition is met, **then** stop

4. **Model adjustment:**
for instance i such that there is y
and $f^*(x_i, y) = 1$ **do**
$z[i] \mathrel{-}= sum[i, y]$
$sum[i, y] \mathrel{\times}= \exp(\alpha^*)$
$z[i] \mathrel{+}= sum[i, y]$

5. **go to** step 1.

Figure 2

0. **Initialize**: S = $\varnothing$, $sum[1..I, 1..Y] = 1$,
$z[1..I] = Y$, $g[1..F] = \{g(1,0),\ldots,g(F,0)\}$

1. **Gain computation:**
MaxGain = 0
Loop
$f_j = \arg\max_{f \text{ in } F}\{g[1,\ldots,|F|\}$

if $g[j] \leq$ MaxGain **then go to** step 2
else
$\hat{\alpha} = \arg\max_{\alpha} G_{S\cup f}(\alpha)$
$\hat{g} = \max_{\alpha} G_{S\cup f}(\alpha)$
$g[j] = \hat{g}$
if MaxGain < $\hat{g}$ **then**
MaxGain = $\hat{g}$
$f^* = f_j$
$\alpha^* = \hat{\alpha}$

2. **Feature selection**:
S = S $\cup$ $\{ f^* \}$
$F = F - \{ f^* \}$

3. **if** termination condition is met, **then** stop

4. **Model adjustment:**
for instance i such that there is y
and $f^*(x_i, y) = 1$ **do**
$z[i]$ $-$= $sum[i, y]$
$sum[i, y]$ $\times$= $\exp(\alpha^*)$
$z[i]$ += $sum[i, y]$

5. **go to** step 1.

Figure 3

Initialization

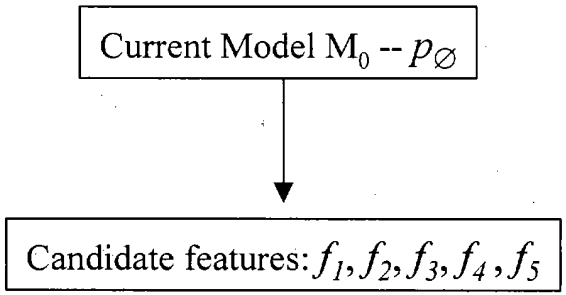

Figure 4

IFS $k=0$

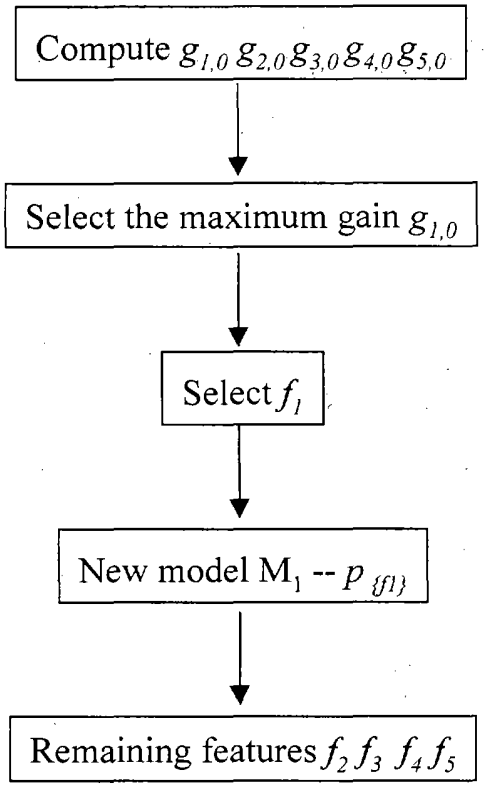

Figure 5A

SGC $k=0$

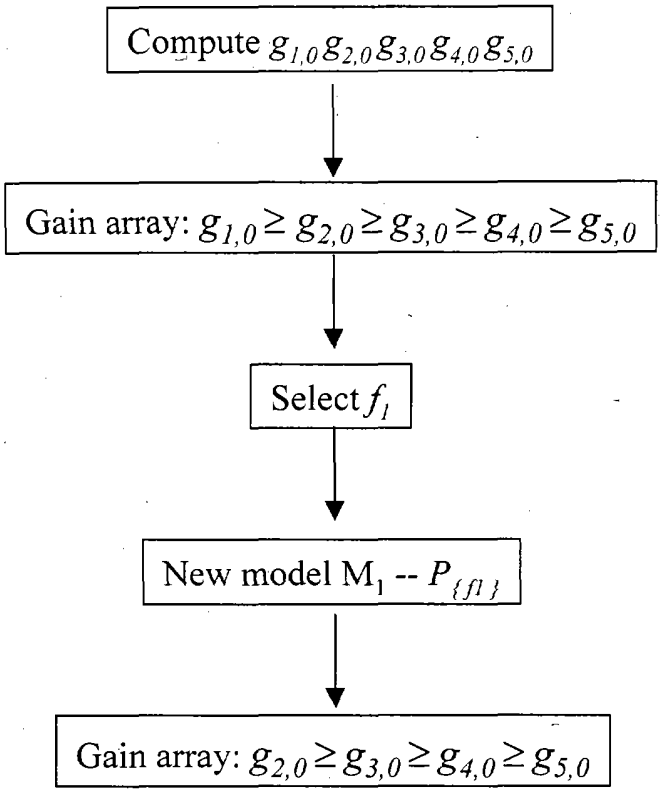

Figure 5B

IFS k=1

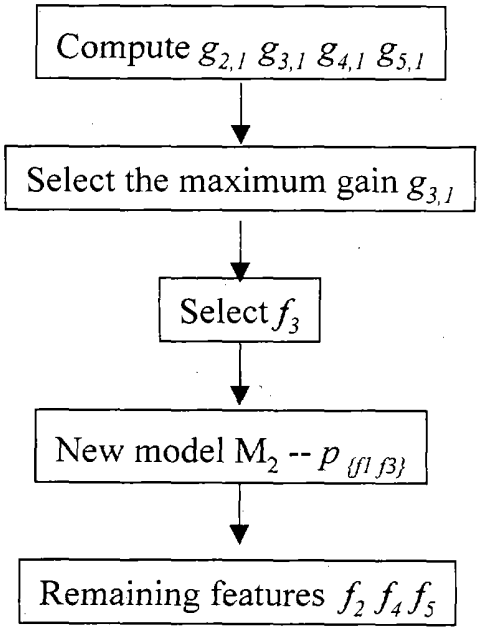

Figure 6A

SGC k=1

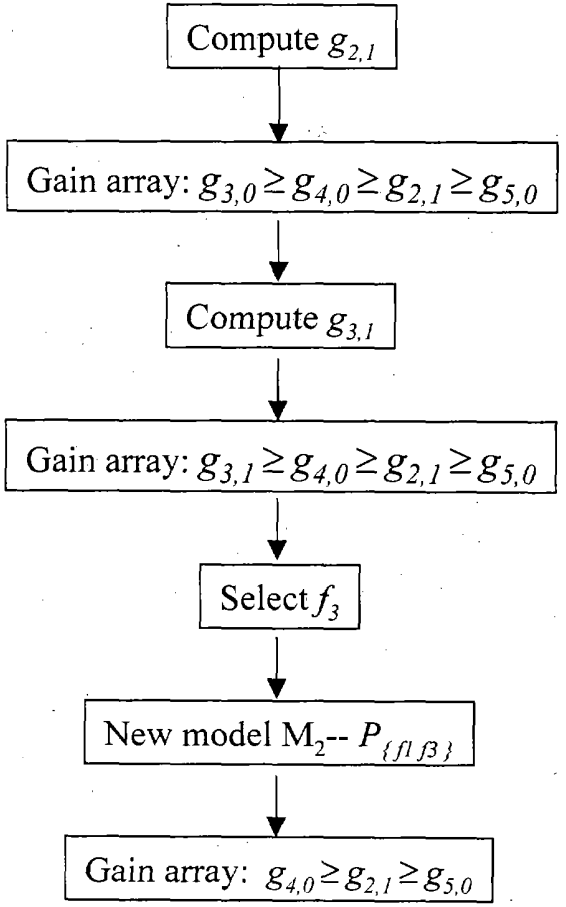

Figure 6B

IFS k=2

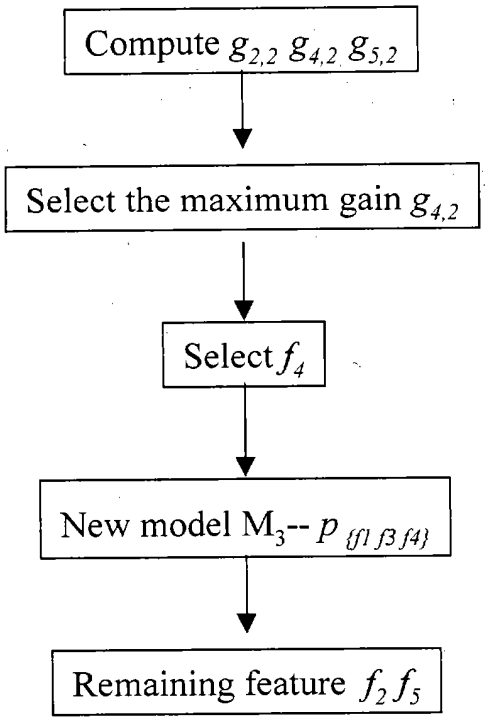

Figure 7A

SGC k=2

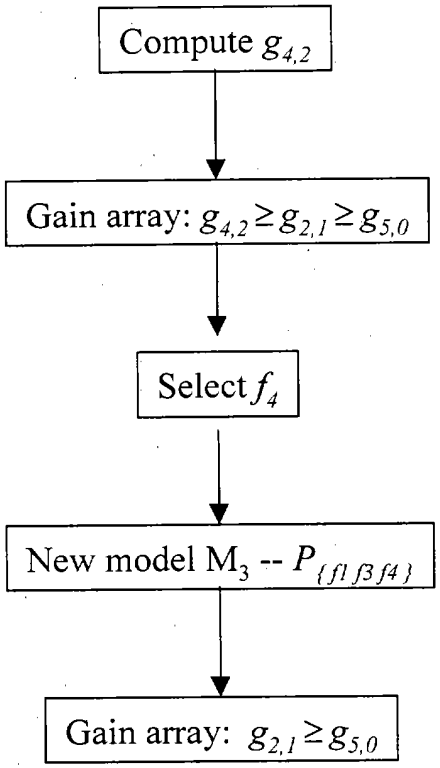

Figure 7B

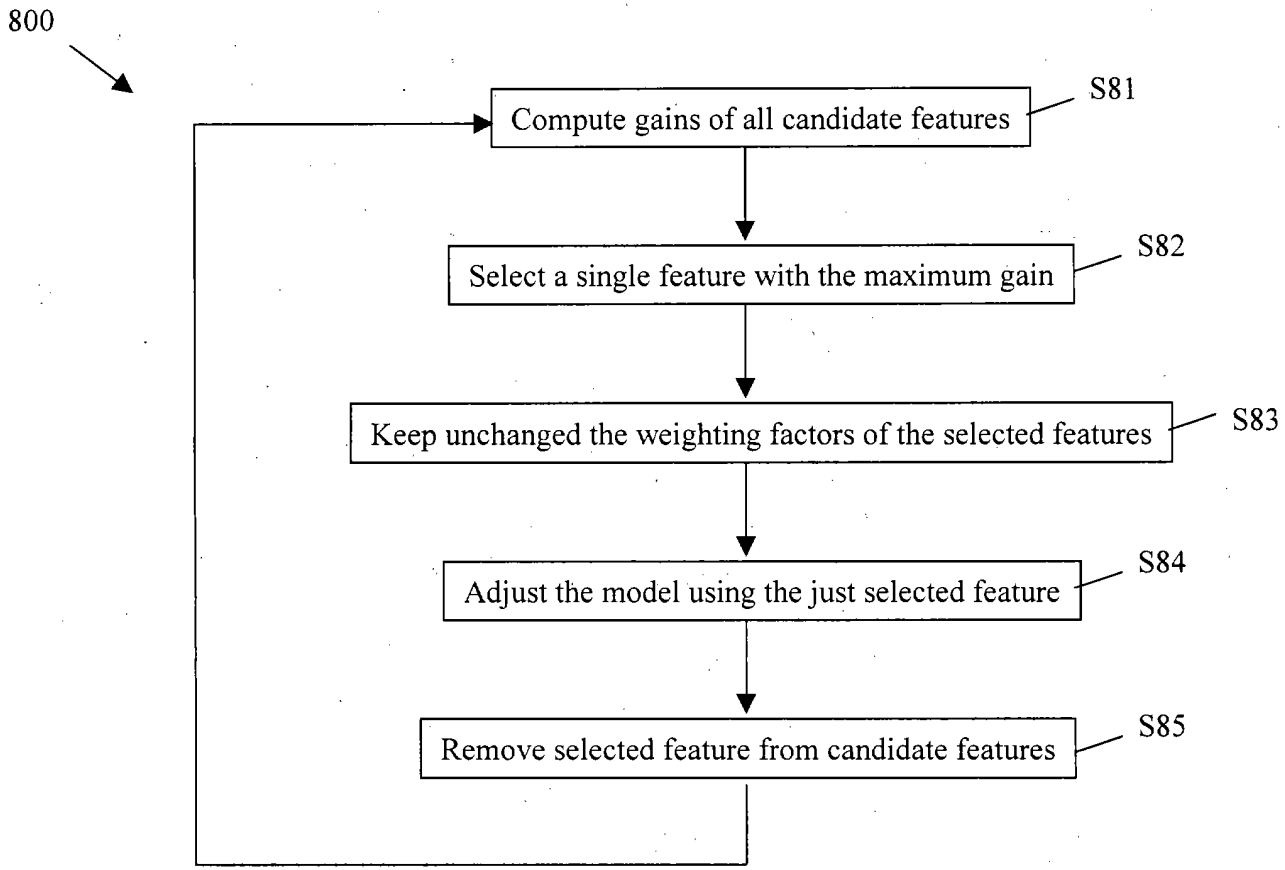
800
Compute gains of all candidate features
S81
Select a single feature with the maximum gain
S82
Keep unchanged the weighting factors of the selected features
S83
Adjust the model using the just selected feature
S84
Remove selected feature from candidate features
S85

Figure 8

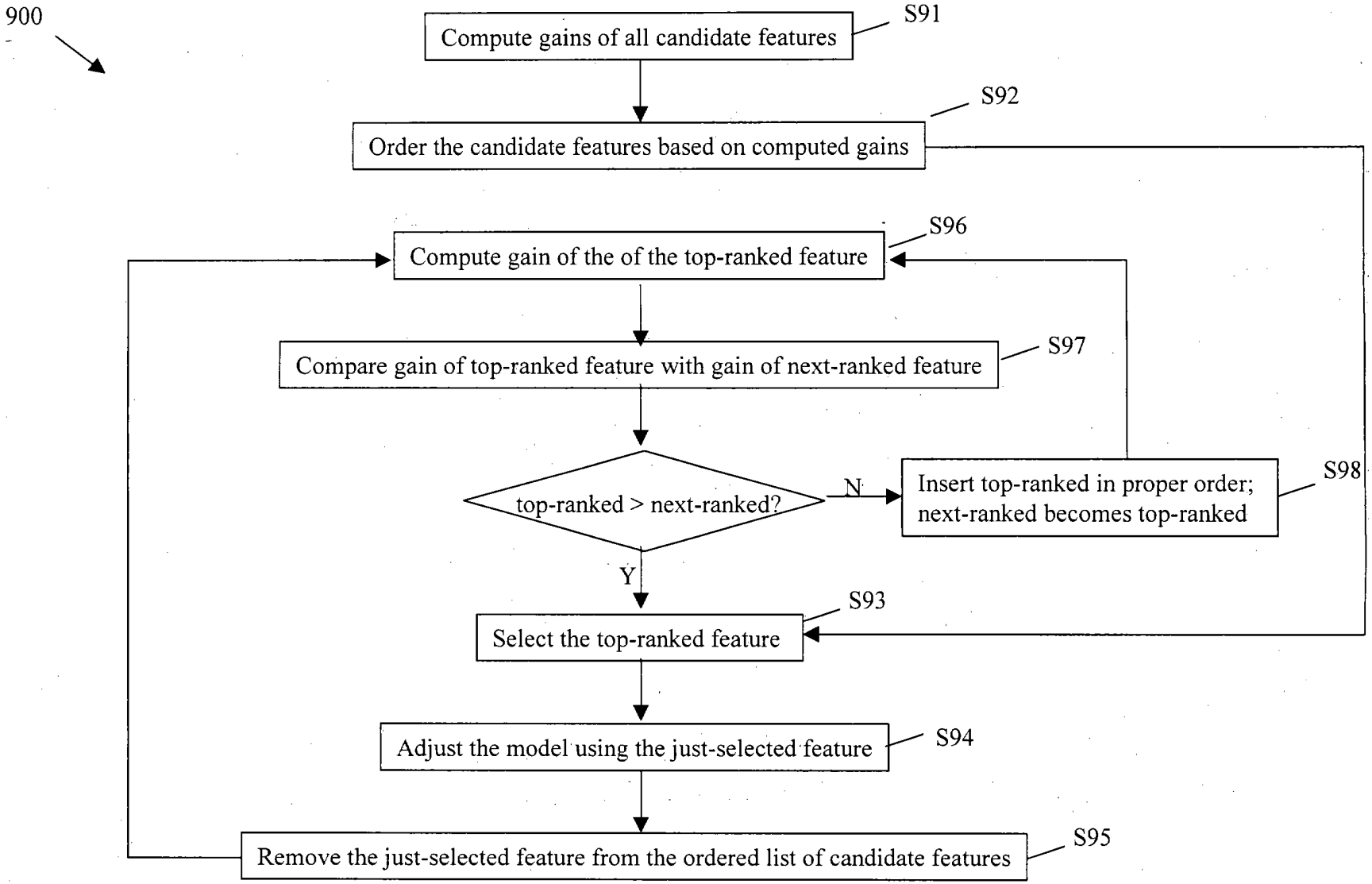
900
Compute gains of all candidate features
S91
Order the candidate features based on computed gains
S92
Compute gain of the of the top-ranked feature
S96
Compare gain of top-ranked feature with gain of next-ranked feature
S97
top-ranked > next-ranked?
N
Insert top-ranked in proper order; next-ranked becomes top-ranked
S98
Y
Select the top-ranked feature
S93
Adjust the model using the just-selected feature
S94
Remove the just-selected feature from the ordered list of candidate features
S95

Figure 9

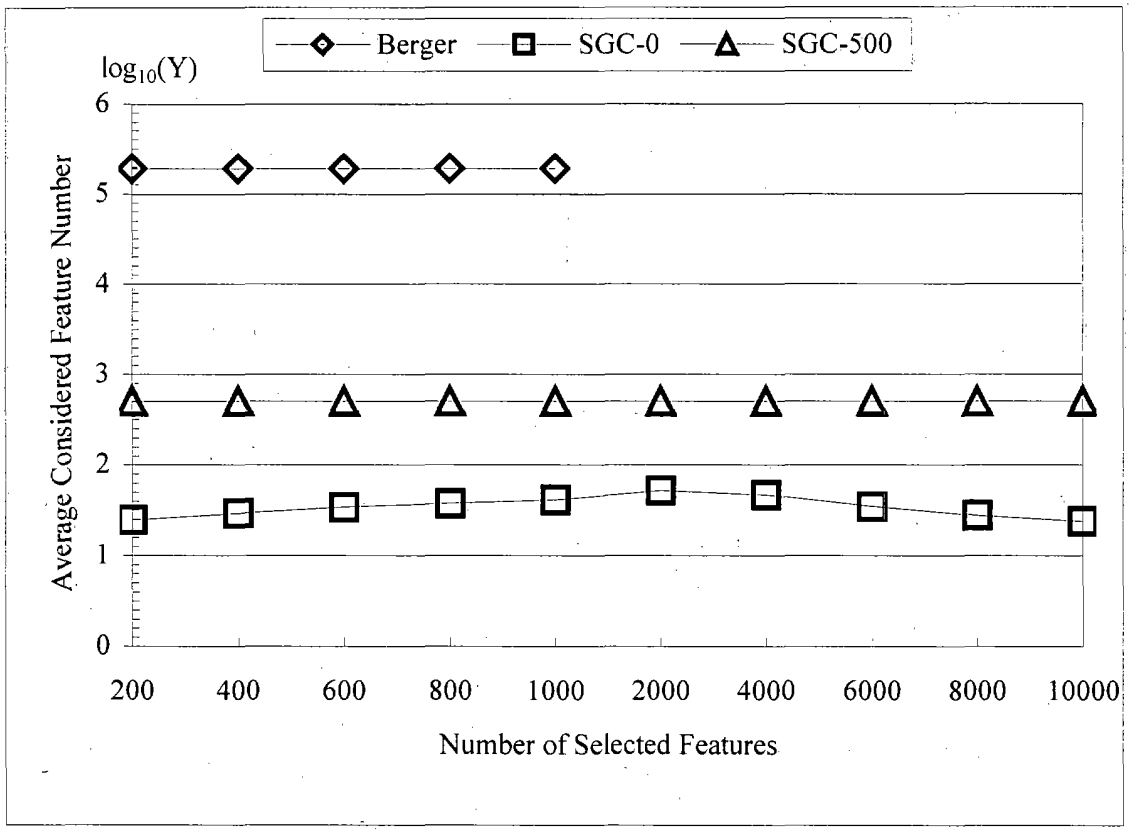
Berger
SGC-0
SGC-500
log10(Y)
Average Considered Feature Number
6
5
4
3
2
1
0
200
400
600
800
1000
2000
4000
6000
8000
10000
Number of Selected Features

Figure 10

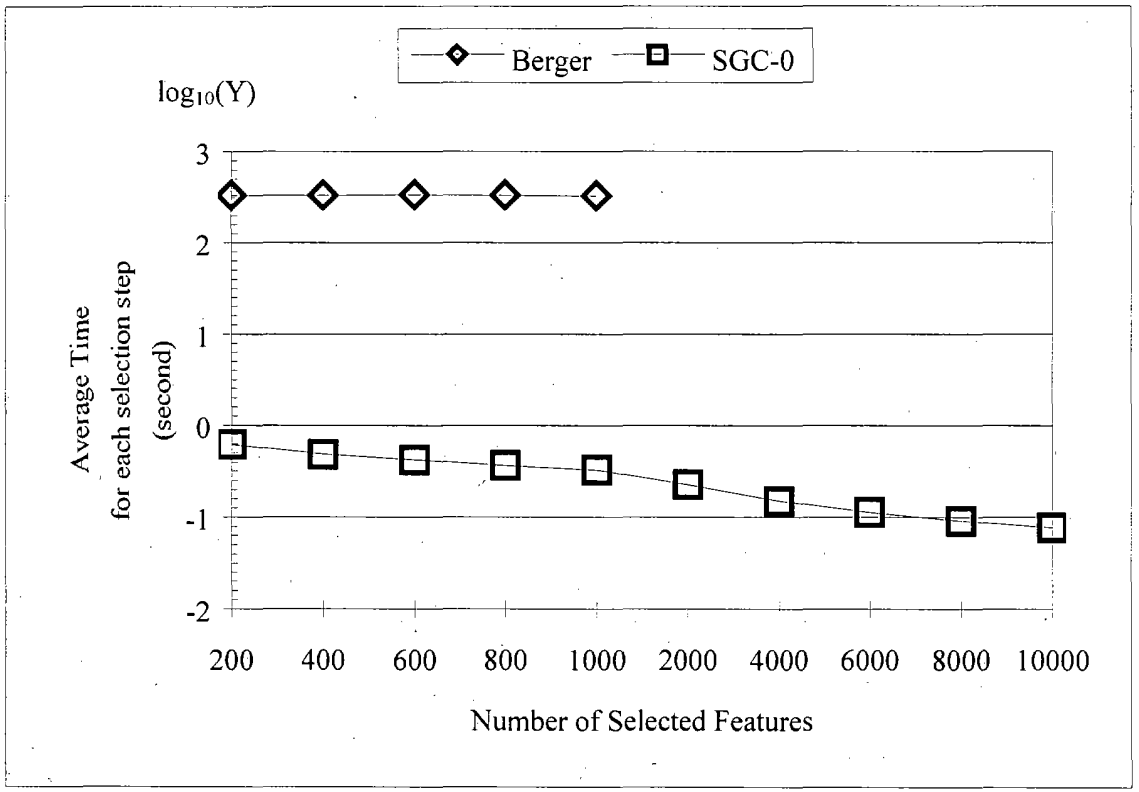
Berger
SGC-0
log10(Y)
3
2
1
0
-1
-2
Average Time for each selection step (second)
200
400
600
800
1000
2000
4000
6000
8000
10000
Number of Selected Features

Figure 11

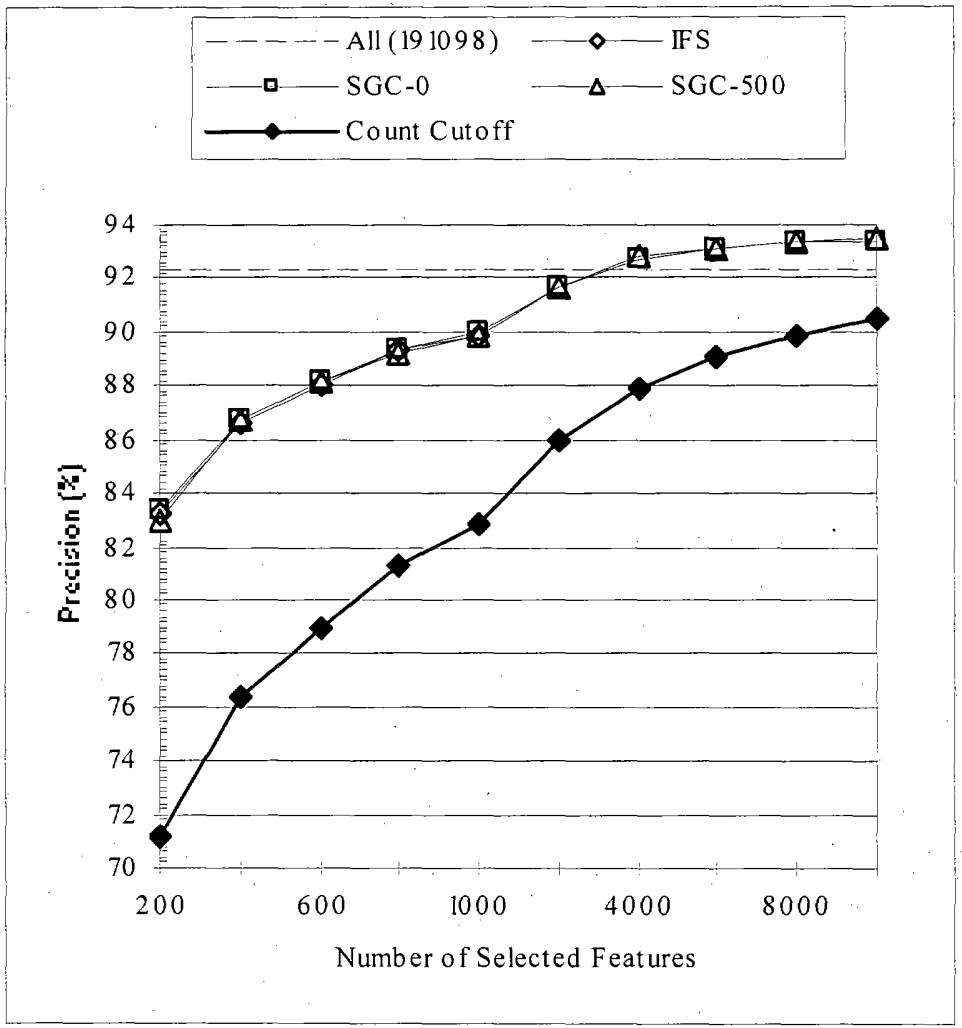
All (191098)
IFS
SGC-0
SGC-500
Count Cutoff
Precision (%)
94
92
90
88
86
84
82
80
78
76
74
72
70
200
600
1000
4000
8000
Number of Selected Features

Figure 12